DOCUMENT RESUME

ED 042 301 EC 006 151

AUTHOR	Staats, Arthur W.; And Others
TITLE	A Token-Reinforcement Remedial Reading Program Administered by Instructional Technicians.
INSTITUTION	Wisconsin Univ., Madison. Research and Development Center for Cognitive Learning.
SPONS AGENCY	Office of Education (DHEW), Washington, D.C.
REPORT NO	TR-127
PUB DATE	May 70
CONTRACT	OEC-5-10-154
NOTE	28p.
EDRS PRICE	EDRS Price MF-$0.25 HC-$1.50
DESCRIPTORS	*Behavior Change, *Disadvantaged Youth, *Exceptional Child Research, Paraprofessional School Personnel, *Reading Achievement, *Reinforcement, Remedial Programs

ABSTRACT

Instructional technicians administered a token reinforcement remedial reading program to 32 black ghetto children who were problem learners. In the 40.2 mean hours of training, the subjects made a mean of 78,505 reading responses; learned a mean of 726.8 words, retaining 81% short term and 59% long term; and received a mean monetary amount of $21.34 worth of reinforcers. Test results indicated that the subjects did significantly better than the controls on a 100-word sample from the training materials. They also excelled on a standard reading achievement test, including the vocabulary section, and on both the verbal and nonverbal portions of an intelligence test. (Author/JD)

ED 042 201

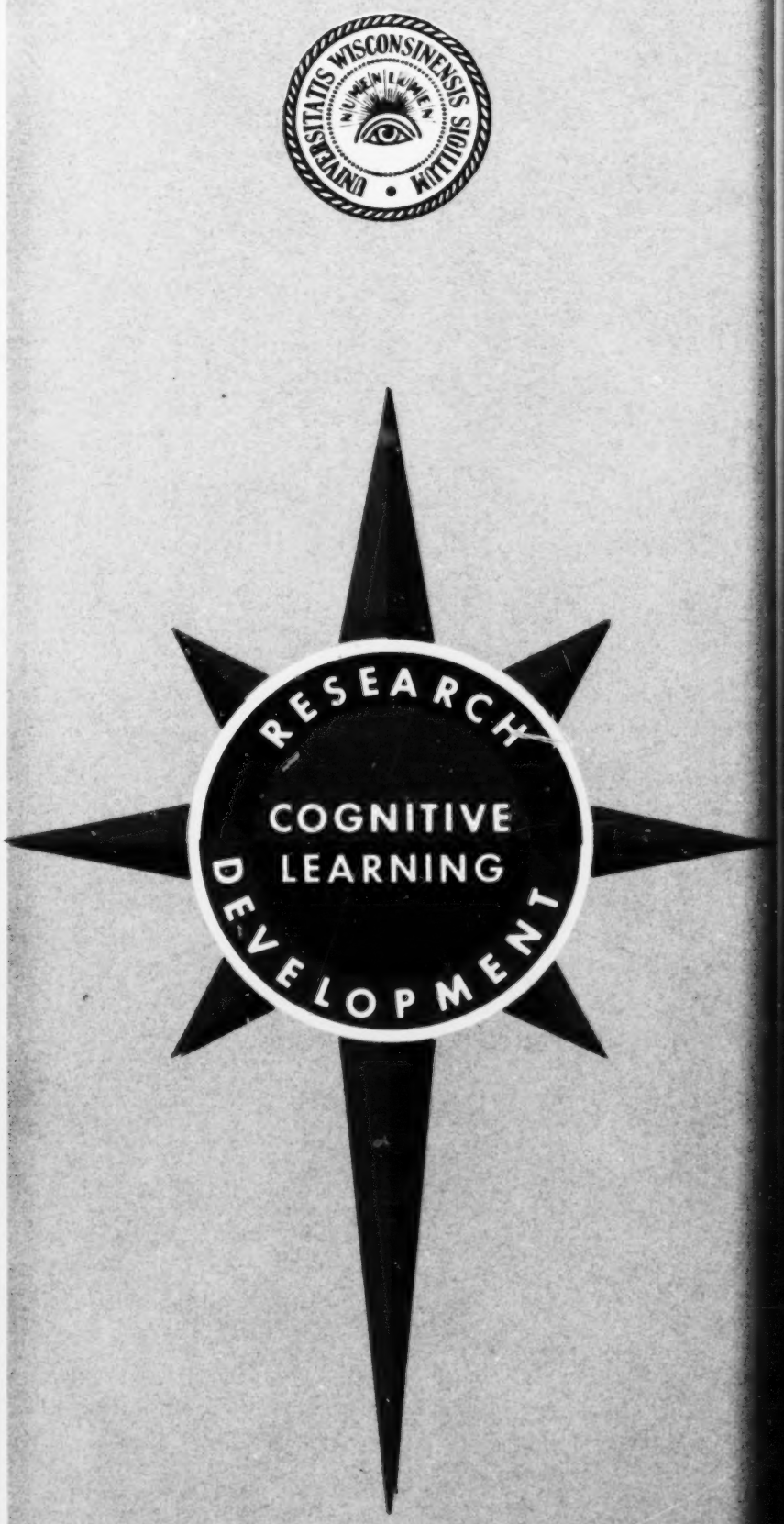
TECHNICAL REPORT NO. 127

A TOKEN-REINFORCEMENT REMEDIAL READING PROGRAM ADMINISTERED BY INSTRUCTIONAL TECHNICIANS

REPORT FROM THE PROJECT ON MOTIVATED LEARNING

WISCONSIN RESEARCH AND DEVELOPMENT

CENTER FOR COGNITIVE LEARNING



ED042301

Technical Report No. 127

A TOKEN-REINFORCEMENT REMEDIAL READING PROGRAM ADMINISTERED BY INSTRUCTIONAL TECHNICIANS

by Arthur W. Staats,* Karl A. Minke,* and Priscilla Butts

Report from the Project on Motivated Learning
Arthur W. Staats, Principal Investigator
*Now in the Department of Psychology,
University of Hawaii

Wisconsin Research and Development
Center for Cognitive Learning
The University of Wisconsin
Madison, Wisconsin

May 1970

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

Published by the Wisconsin Research and Development Center for Cognitive Learning, supported in part as a research and development center by funds from the United States Office of Education, Department of Health, Education, and Welfare. The opinions expressed herein do not necessarily reflect the position or policy of the Office of Education and no official endorsement by the Office of Education should be inferred.

Preparation of this manuscript was in part supported by the Educational Research and Development Center of the University of Hawaii.

Center No. C-03 / Contract OE 5-10-154

NATIONAL EVALUATION COMMITTEE

Samuel Brownell
Professor of Urban Education
Graduate School
Yale University

Launor F. Carter
Senior Vice President on Technology and Development
System Development Corporation

Francis S. Chase
Professor
Department of Education
University of Chicago

Henry Chauncey
President
Educational Testing Service

Martin Deutsch
Director, Institute for Developmental Studies
New York Medical College

Jack Edling
Director, Teaching Research Division
Oregon State System of Higher Education

Elizabeth Koontz
Wage and Labor Standards Administration, U.S. Department of Labor, Washington

Roderick McPhee
President
Punahou School, Honolulu

G. Wesley Sowards
Director, Elementary Education
Florida State University

Patrick Suppes
Professor
Department of Mathematics
Stanford University

***Benton J. Underwood**
Professor
Department of Psychology
Northwestern University

RESEARCH AND DEVELOPMENT CENTER POLICY REVIEW BOARD

Leonard Berkowitz
Chairman
Department of Psychology

Archie A. Buchmiller
Deputy State Superintendent
Department of Public Instruction

Robert E. Grinder
Chairman
Department of Educational Psychology

Russell J. Hosler
Professor, Curriculum and Instruction

Clauston Jenkins
Assistant Director
Coordinating Committee for Higher Education

Herbert J. Klausmeier
Director, R & D Center
Professor of Educational Psychology

Stephen C. Kleene
Dean, College of Letters and Science

Donald J. McCarty
Dean
School of Education

Ira Sharkansky
Associate Professor of Political Science

B. Robert Tabachnick
Chairman, Department of Curriculum and Instruction

Henry C. Weinlick
Executive Secretary
Wisconsin Education Association

M. Crawford Young
Associate Dean
The Graduate School

EXECUTIVE COMMITTEE

Edgar F. Borgatta
Brittingham Professor of Sociology

Anne E. Buchanan
Project Specialist
R & D Center

Robin S. Chapman
Research Associate
R & D Center

Robert E. Davidson
Assistant Professor, Educational Psychology

Frank H. Farley
Associate Professor, Educational Psychology

Russell J. Hosler
Professor of Curriculum and Instruction and of Business

***Herbert J. Klausmeier**
Director, R & D Center
Professor of Educational Psychology

Wayne Otto
Professor of Curriculum and Instruction (Reading)

Robert G. Petzold
Associate Dean of the School of Education
Professor of Curriculum and Instruction and of Music

FACULTY OF PRINCIPAL INVESTIGATORS

Vernon L. Allen
Professor of Psychology

Ted Czajkowski
Assistant Professor of Curriculum and Instruction

Robert E. Davidson
Assistant Professor of Educational Psychology

Gary A. Davis
Associate Professor of Educational Psychology

M. Vere DeVault
Professor of Curriculum and Instruction (Mathematics)

Frank H. Farley
Associate Professor of Educational Psychology

Lester S. Golub
Lecturer in Curriculum and Instruction and in English

John G. Harvey
Associate Professor of Mathematics and of Curriculum and Instruction

Herbert J. Klausmeier
Director, R & D Center
Professor of Educational Psychology

Donald Lange
Assistant Professor of Curriculum and Instruction

James Moser
Assistant Professor of Mathematics Education; Visiting Scholar

Wayne Otto
Professor of Curriculum and Instruction (Reading)

Milton O. Pella
Professor of Curriculum and Instruction (Science)

Thomas A. Romberg
Associate Director, R & D Center
Professor of Mathematics and of Curriculum and Instruction

B. Robert Tabachnick
Chairman, Department of Curriculum and Instruction

Richard L. Venezky
Assistant Professor of English and of Computer Sciences

Alan Voelker
Assistant Professor of Curriculum and Instruction

Larry Wilder
Assistant Professor of Curriculum and Instruction

Peter Wolff
Assistant Professor of Educational Psychology

MANAGEMENT COUNCIL

Herbert J. Klausmeier
Director, R & D Center
V.A.C. Henmon Professor of Educational Psychology

Mary R. Quilling
Director
Technical Development Program

Thomas A. Romberg
Associate Director

James Walter
Director
Dissemination Program

Dan G. Woolpert
Director
Operations and Business

* COMMITTEE CHAIRMAN

STATEMENT OF FOCUS

The Wisconsin Research and Development Center for Cognitive learning focuses on contributing to a better understanding of cognitive learning by children and youth and to the improvement of related educational practices. The strategy for research and development is comprehensive. It includes basic research to generate new knowledge about the conditions and processes of learning and about the processes of instruction, and the subsequent development of research-based instructional materials, many of which are designed for use by teachers and others for use by students. These materials are tested and refined in school settings. Throughout these operations behavioral scientists, curriculum experts, academic scholars, and school people interact, insuring that the results of Center activities are based soundly on knowledge of subject matter and cognitive learning and that they are applied to the improvement of educational practice.

Activities are differentiated into three research and development Programs —1, Conditions and Processes of Learning; 2, Processes and Programs of Instruction; and 3, Facilitative Environments—and support programs. This Technical Report is from Project Motivated Learning in Program 2. General objectives of the Program are to generate new knowledge about concept learning and cognitive skills, to synthesize existing knowledge, and to develop materials suggested by the prior activities. Contributing to these Program objectives, this project aims at extending learning theory in the context of complex human behavior, and applying the theory in developing procedures to solve problems of human learning.

CONTENTS

		Page
	Abstract	ix
I	Introduction	1
II	Method	5
	Subjects	
	Experimenters (Therapy Technicians or Instructional Technicians)	6
	Instruments	7
	The Token Reinforcer System for the Behavior Modification Procedure	7
	Reading Materials	7
	Vocabulary Words	7
	Oral Reading Materials	8
	Procedure	8
	Vocabulary Presentation	8
	Oral Reading	8
	Silent Reading and Comprehensive Questions	8
	Vocabulary Review	8
	Facilities	9
III	Results and Conclusions	11
	Behavioral Data	11
	Test Data	14
IV	Discussion	17
	References	23

LIST OF FIGURES AND TABLE

Figure		Page
1	Cumulative number of reading responses as a function of the time in reading training for Ss enrolled in EMR and regular classes	12
2	The ratio of words missed upon initial presentation to the total number of words presented in the Individual Word Phase or a function of the number of SRA stories read for the two classes	13
3	Ratio of the monetary value of the tokens received divided by the number of reading responses mode as a function of time in reading	13
4	Mean scholastic grade point average in academic subjects for Experimental and Control Ss across the two semesters of the school year	14

Table		
1	Analysis of Covariance on Pre- and Post-Test Measures	15

ABSTRACT

Reading materials in a stimulus-response presentation procedure combined with a token-reinforcement motivational system were employed with 32 ghetto black children in a remedial program. The program was administered by black subprofessional instructional technicians supervised by a teacher trained in the experimental methods. The children, who were problem learners—many considered to be emotionally disturbed, anti-social, or retarded—attended well, worked hard, and learned well in the 4- to 5-month program. There were two types of dependent measures: behavioral measures which consisted of detailed recording of the responses each child made, the reinforcers received, the words learned and retained, and the like, and also test results. The 32 experimental subjects made a mean of 78,505 reading responses in the mean 40.2 hours of training. They learned a mean of 726.8 words and retained 81% in a short-term test and 59% in a long-term test. The experimental subjects received a mean of $21.34 worth of reinforcers. On the test data the experimental group showed significant increase in their ability to read a 100-word sample of the 4,253 different words included in the training materials, as compared to a matched group of 32 control subjects. On a standard reading achievement test the experimental subjects were significantly better (at the .07 level) than the control group on the reading portion, which was most similar to their training materials. On the vocabulary portion, which was composed of very dissimilar items from the training materials, the experimental group improved more than the controls, but only to a .14 level of reliability. On an intelligence test the experimental group improved more than the controls on the non-verbal portion (significant at the .04 level)—which involved items like material of the experimental training—as well as on the verbal portion (significant at the .12 level). The results showed the procedures and reinforcement system to be effective in producing improved attention and work behaviors in these usually intractable children, and in the utilization and upgrading of unemployed black adults. However, 40.2 hours of training were not deemed sufficient to remediate long-standing cases of educational failure, although the results suggested that longer, more intensive, programs of the present kind could make important contributions to the social problems involved as well as to the study of human learning.

I
INTRODUCTION

A number of years ago as a student at UCLA, the first author[1] received experience in the late Grace Fernald's clinical school for children with learning difficulties. He concluded at that time that a major problem with the children's learning was their lack of attention and poor work behavior because of inadequacies in reinforcement (motivation). The study of learning and behavior problems in children remained one of his focal interests and later became part of an extended project of research which he has conducted.

By 1959 he had developed a token-reinforcer system for working with backward children with reading problems. That is, the children were given plastic discs for reading (in the type of program to be described). These discs could be exchanged for rewards the children had chosen to work for. Under such a token-reinforcement system the children metamorphosed into attentive, hardworking, students. This was the first development of the token-reinforcement system for behavior modification studies. The system was then used by John Michael at the University of Houston in work with retardates and by his students, Patricia Cooke and Sam Toombs, in a classroom for remedial reading. The token-reinforcement system has since been widely employed in various forms in educational and clinical behavior modification studies (for example, see Wolf, Giles, & Hall, 1968, and Ullman & Krasner, 1965). After the initial development of the token-reinforcement system the first author also adapted it for work with preschool children in a series of basic and behavior modification studies (See Staats, 1968a, 1968c; Staats, Staats, Schutz, & Wolf, 1962; Staats, Finley, Minke, Wolf, & Brooks, 1964; and Staats, Minke, Finley, & Wolf, 1964.)

These studies involved the investigation of original number concept learning, writing, and especially reading. However, in continuing the study of cognitive learning in children with special problems, the token-reinforcement system was employed in additional work. In one study a culturally deprived juvenile delinquent child with severe learning and behavior problems was successfully treated with the token-reinforcement system (Staats & Butterfield, 1965). This study also involved stimulus materials for reading training which the first author derived from a theoretical-experimental analysis of the learning task. As a consequence of the explicit analysis, the instructional task (and data recording) could be designed to be simple and straightforward. Administration of the materials thus did not require a trained teacher. This suggested that subprofessional instructional technicians could be employed in the remedial reading treatment. A study was then conducted with 18 additional children in Madison, Wisconsin, to test this possibility. The instructional technicians were volunteer housewives and high school students. Again, the efficacy of the token-reinforcement system and the stimulus materials and procedures was validated (Staats, Minke, Goodwin, & Landeen, 1967).

---

[1]The first author planned the experiment and wrote the present paper as principal investigator of Project Motivated Learning under the support of the University of Wisconsin Center for Cognitive Learning (OE Center No. C-03). The second author was responsible for the data recording, the statistical analysis of the tests, and assisted in collating the results and describing them. The third author supervised the instructional technicians and trained them also, with the help of the second author.

One of the primary social problems in our society is that of the general failure of education with culturally deprived children. In a far greater proportion than other segments of our society, children from economically depressed, ghetto homes do poorly in school and either drop out or ultimately leave school without having acquired the cognitive skills necessary for successful competition in our society. Poorly trained individuals are unable to get high-status, good-paying positions, even when discrimination on the basis of color is not a problem. Such individuals, then, are not in a position to raise children who have a better chance at advancement.

Methods are needed for penetrating this vicious cycle and improving the educational advancement of culturally disadvantaged people, both children and young adults. As indicated, the application of reinforcement principles to the treatment of individual problems of behavior modification in clinical psychology is becoming widespread, with increasing general acceptance. It has been suggested that the same principles should also be extended extensively and systematically to group and social problems (Staats, 1968a, 1970). In the present case the social problem is that of the culturally disadvantaged child in schools designed for children with different training backgrounds. Such disparities between the child and the school result in the failure of both parties.

Specifically, it is again suggested that motivational problems largely underlie educational failure including that which is disproportionately evidenced among ghetto Black youth. In the Staats and Butterfield study (1965), for example, the case of the culturally disadvantaged, educationally retarded, juvenile delinquent is described in terms of such motivational principles. It was suggested that the reward value of learning, of grades, of competition for grades with other children, of social approval from white adults, of familial social approval for school achievement, and so on, may be entirely lacking for the culturally disadvantaged child. The "reinforcement system" of the school, however, is built around such rewards. This means there is an absence of effective rewards for maintaining the long years of attention, participation, and work involved in becoming academically and cognitively accomplished. A high percentage of failure must be expected in cases where the reinforcement system of the institution is inappropriate for the learned reinforcement system of the individual and his group (Staats, 1968a, 1970). The study supported the thesis that motivation; that is, the principles of reinforcement, are at the heart of many group and social problems in education.

The present study extends these findings to the social problem of the failure of education with the ghetto black child. The aim of the study is to experimentally test several expectations:

(1) A primary concern is to test the possibility that the attentional and work behaviors of culturally disadvantaged children who have learning problems can be improved through the use of the reward and training procedures developed earlier.

(2) The study also attempts to explore the expectation that such children will learn well and retain academic material when it has been learned. Methods are utilized in the study to record the specific learning and retention of the subjects. It is thus possible to indicate the number of reading responses made, the number of new words learned, the extent to which these words are retained, and so on. These behavioral data are the major types of data of the study.

(3) The previous studies in the project have employed adaptations of Science Research Associates' reading materials, and the present study assesses the effects of training with such materials on reading achievement tests, the other types of data in the study.

(4) As suggested, when a cognitive learning task is analyzed into its stimulus response constituents, one should be able to construct explicit procedures for producing the learning. In the present case it was shown that subprofessional technicians could administer the training program. Actual application of this possibility on a wider scale demands an organizational structure, as well as a set of training materials and procedures. The plan was to train a teacher in the remedial methods employing the token-reinforcer system. This teacher (the third author) then selected, trained, and supervised a team of instructional technicians who actually administered the reading training. The instructional technicians were also to be /black/ (as the teacher could have been). Two types of instructional technicians were used: formerly unemployed, literate, adults were employed full-time; and adult volunteers, as well as literate high school seniors in ghetto schools, were used to work with individual children. The full-time instructional

technicians were used to assess the possibility of solving problems of unemployment among black adults as well as the problem of upgrading employment conditions for Black people at the same time that the educational problems were being treated.

(5) The study is also a test of the learning theory underlying the procedures. Although the basic principles of reinforcement have been precisely isolated in the laboratory, they only serve as basic principles for a theory of human behavior when they are verified in the context of complex human behavior. The present type of study, in being concerned with the verification of basic learning principles, serves a role in the development of the human learning theory, which is as central as the laboratory isolation of the principles. As such, the study is thus relevant to both applied and basic fields.

(6) At the present time most behavior modification studies have involved single case studies (see Ullman & Krasner, 1965). Thus and last, the present study begins the attempt to apply a behavior modification procedure to a large enough sample of subjects to verify the procedures by replication. The behavioral results of the use of the reinforcement system and stimulus materials with 32 subjects could provide such replication data. In addition, a control group was used so that statistical comparisons on some behavioral measures (such as grades and attendance) could be made along with comparisons on standardized tests. The tests were included to begin the assessment of the generality of the training effects.

II
METHOD

SUBJECTS

Twelve experimental and 12 control Ss were selected from each of three junior high schools servicing the Inner Core area of Milwaukee, Wisconsin. All Ss were of Afro-American heritage. Eight of the 24 Ss at each school were enrolled in special classes for the educably mentally retarded (EMR), and the remaining 16 Ss were enrolled in regular classes at the Seventh, Eighth, or Ninth Grade level. The mean age of the experimental subjects at the beginning of the study was 14 years and 2 months. The mean intelligence age of the children, as given on the Lorge-Thorndike Intelligence Test, Level C, in an age-equivalent IQ score, was 8.26. This age equivalent score was in all likelihood an over-estimate since the form of the test employed was appropriate for Third and Fourth Graders and would thus have included easier elements than the higher level test appropriate for this age group. At any rate, as tested, these children with a mean age of 14 years and 2 months were functioning on the test as children whose age was 8.26 years. For the EMR children the mean age equivalent score was 7.43; for the children in regular classes, the mean score was 8.70. The mean chronological age of the control subjects was 14 years and 1.2 months. Their mean age equivalent IQ scores were 8.59. Lists of students defined as poor readers by their teachers and counselors were obtained from each type of class at each of the three schools. Final selection of Ss for participation in the study was done on the basis of a 100-item word recognition test, a sample of words taken from the reading material. The reading materials were for Grade 1 to Grade 4 levels.

The 24 Ss from special educational classes and the 48 Ss enrolled in regular classes were selected randomly from among those students in each type of class at each of the three schools who read no more than 60 of the items contained in the word recognition test nor less than 10. The Ss so selected were divided into 36 pairs, matched on the basis of their scores on the word recognition test, the particular school they attended, and the type of class in which they were enrolled (EMR or regular). The 36 Ss for the experimental group were obtained by randomly selecting one of the two students from each matched pair.

Four experimental Ss were lost to the project during the course of the experiment (one S was not reinstated following suspension from school and three Ss were sent to the reformatory, the latter including two of the three children who were from a special class for returnees from the reformatory). Thus the data reported herein are based upon 32 experimental and 32 of the control Ss. Of the 32 experimental Ss, 19 were identified as severe behavior problems, and a number had spent time in the detention home or reform school. Some control Ss were also lost during the project but were replaced from a pool of alternates who were given the various pretests at the beginning of the experiment; a replaced control S was matched with his experimental S on the basis of the same criteria used in the original matching.

Ss were identified as culturally disadvantaged on a geographical basis; that is, they resided within the Inner Core area of the city, as well as on the basis of their cognitive deficits. Information collected at the end of the experiment for the experimental Ss supported the contention that the Ss were a sample from a culturally disadvantaged population. Eight of the families were on relief, and in 20 of the remaining 24 families, the wage earner worked as an unskilled laborer. The remaining four families were supported by a student, store owner, a clerk, and an industrial foreman. Nine of the families were being supported solely by the mother. The average income per family was $5,109 with a mean number of 7.56 persons in each family. A frequency

distribution of average family income showed four families had a yearly income between $1,500 and $3,000 or 10 families had an income between $3,000 and $4,500 or eight families were in the next interval of $4,500-$6,000 and five families had incomes in both the $6,000-$7,500 and $7,500-$9,000 brackets.

The mean number of school years completed by the mother was 9.34, and 7.86 by the father. It had been suggested by the school administration that the majority of the Ss would be drawn from a group of families who had recently migrated to Wisconsin from southern states, since it was their observation that these children were in general more backward and intransigent than children who had spent the majority of their school years in the Milwaukee school system. However, the experimental families had lived in Milwaukee for an average of 14 years, and 15 of the 32 children had been born in Milwaukee. The subjects randomly assigned to the control group were given the same tests as were the experimental group subjects. The control subjects, however, received no further special treatment, only attending school as usual.

EXPERIMENTERS (THERAPY TECHNICIANS OR INSTRUCTIONAL TECHNICIANS)

Three types of instructional technicians were utilized in this study, each type being used at a different school. Six adult volunteers were obtained through a list of the applicants for the positions of social welfare aide and Headstart aide in the ghetto community. Each of these volunteers worked with two children and were paid $1.85 a session to cover transportation, baby-sitting fees, etc. The adult volunteers were all housewives from the Inner Core. Their average age was 37, and each tutor had an average of 3.5 children living at home with an average age of 10.5 years. Instructional technicians in this group had a mean family income of $6,500 and had completed an average of 12.25 years of school. Two of the six tutors had employment in addition to working in the project. All potential instructional technicians were administered the reading portion of the Wide Range Achievement Test (1965 edition), and all individuals with a score below 8th Grade reading ability were eliminated from consideration. The mean grade level of this group of part-time volunteers as measured on this instrument was 10.3.

At a second school two housewives were hired on a full-time basis at $360 per month. Each of these instructional technicians worked with six children across the day and used their additional time to assist in data collection and collation. Because these two women would be at the school on a full-time basis and thus be unofficial members of the faculty, the principal insisted upon the final selection of these instructional technicians himself. As a consequence, although both tutors were residents of the Inner Core area, they were not entirely representative of the ghetto population. Their average age was 35.5 years. One instructional technician had no children, while the other tutor had six children living at home, with an average age of 15 years. The mean family income of these two tutors was $9,000. It should be noted, however, that a large portion of this income was a function of the tutor's earning power, since the husbands of both instructional technicians were employed as laborers. Both full-time instructional technicians had completed a year at a small Negro college in Arkansas. The mean grade level on the reading portion of the Wide Range Achievement Test for these two instructional technicians was 12.3 years.

Finally, 12 high school students in either their junior or senior year were employed as instructional technicians at $1.25 an hour. These students each worked with one S during a regularly scheduled study hall. The average age of these 12 instructional technicians was 16.5 years and the average annual income of their families was $5,375. Three of the 12 therapy technicians had, in addition to their participation in this program, after-school jobs. These students had a mean grade average of 2.58 (out of a 4.0 possible) for their high school careers. The mean grade level on the reading portion of the Wide Range Achievement Test for this group was 10.03.

All of the instructional technicians lived within the geographic boundaries of the Inner Core, and all of the adult administrators were black. Two of the high-school therapy technicians were white, however, and two were Puerto Rican. The instructional technicians were trained in the use of the procedures in two sessions, each approximately 1 1/2 hours long. The majority of the second session was spent in actually administering the materials.

The instructional technicians were supervised throughout the project by a teacher (the third author) trained and experienced in the present methods. She observed each tutor on his first day with his Ss and she also periodically observed an experimental session with each of the instructional technicians. The teacher also took the responsibility of administering the day-to-day aspects of the research project.

INSTRUMENTS

The primary test used, both for grouping purposes and as a dependent variable, consisted of 100 words randomly selected from the reading materials used in the experimental sessions with 20 words selected from each of five grade levels (1.2, 1.7, 2.3, 3.0, and 4.0). Subjects had to pronounce correctly the word shown them on a flash card in order to receive credit. The same words comprised both the pretest and posttest forms of this instrument. It should be noted that 4,253 different words were involved in the reading materials. The 100-word test is thus a representative sample of that universe of words. It may help to characterize this sample by indicating that 21 of the 100 words appear in the Dolch basic sight vocabulary (Dolch, 1950).

In addition, two forms of the reading and vocabulary portions of the Metropolitan Achievement Tests for Grades 3 and 4 and two forms of the verbal and non-verbal batteries of the Lorge-Thorndike Intelligence Test, Level C, were utilized, one form of each test being used as a pretreatment measure and the other form as a posttest. In an earlier study (Staats, Minke, Goodwin, & Landeen, 1967) it was found that treatment differences were not obtained between pre- and post-test measures on standardized tests, and it was suggested that one possible reason was that although the tests were standardized with respect to the grade level of the population from which the experimental Ss were drawn, the tests might have been too difficult for the particular children used in the experiment; that is, those extremely backward in reading. Thus, in the present experiment forms of both the achievement test and the intelligence test were utilized which were designed for the elementary school level rather than the junior high school level of the Ss.

THE TOKEN REINFORCER SYSTEM FOR THE BEHAVIOR MODIFICATION PROCEDURE

There were three types of tokens, distinguished by color. The tokens were of different value in terms of the items for which the tokens could be exchanged. A blue token was valued at 1/10 of one cent, a yellow token at 1/5 of a cent, and a red token at 1/2 of one cent.

After S had completed the first 100 lessons of the program, a bonus system was introduced in order to maintain a $.20 minimum in terms of S's daily earnings. From this point on, whenever S's earnings fell below $.20 for any one session, he was given a bonus of $.05, $.10, or $.15, whichever amount was needed in order to bring his earnings for that day up to between $.20 and $.25. When the bonus system was introduced, the child was told that he could not move as rapidly as before because the material upon which he was working was difficult, but if he continued to work hard he would be given a bonus at the end of the session. This bonus system was introduced prior to the completion of 100 lessons for several extremely slow Ss.

The child's acquisition of tokens was plotted so that visual evidence of the reinforcers was available. The tokens could be used to purchase a variety of items. These items, chosen by the subject, could range in value from pennies to whatever the subject wished to work for. Records were kept of the tokens earned by S and of the manner in which the tokens were used.

READING MATERIALS

The reading materials were the same as were used in the previous study (Staats, Minke, Goodwin, & Landeen, 1967) and were taken from the Science Research Associates (SRA) reading-kit materials, Reading Laboratories IA, IB, IC, and IIA. The SRA kits consist of stories developed for and grouped into grade levels. For the purposes of this study, stories were taken from the 1.2 (20 stories), 1.4 (60 stories), 1.7 (20 stories), 2.0 (32 stories), 2.3 (12 stories), 2.6 (12 stories), 3.0 (32 stories), 3.5 (32 stories), and 4.0 (10 stories) grade levels. Once a particular Reading Laboratory was selected for inclusion at a given grade level, all the stories at that grade level were presented in sequential order (with the exception of the 4.0 grade level, where only the first 10 stories in Laboratory IIA were presented). The different numbers of lessons presented at each grade level were due to the use of different numbers of Reading Laboratories, in order to control somewhat the rate of introduction of new words.

The reading training program which was that of the earlier studies (Staats & Butterfield, 1965; Staats, Minke, Goodwin, & Landeen, 1967), may be summarized as follows.

Vocabulary Words

A running list was made of the new words that appeared in the series of stories. Each different form of a word was counted as a different word for this purpose; thus, bring,

brings, and bringing were all counted as different words. The list finally included each different word that appeared in the stories, a total of 4,253 words. From this list, the new vocabulary for each story was selected, and each word was typed on a separate 3" x 5" card.

The average number of new words introduced in each story was 18.5, the least number of new words being introduced in any one story being 6, and the most being 50. The new stories presenting over 30 new words occurred in the second half of the program, after the children had already been presented a large number of lessons.

Oral Reading Materials

Each paragraph in the SRA stories was typed on a 5" x 8" card. Each story could thus be presented to S paragraph by paragraph.

Silent-reading and comprehensive-question materials. Each SRA story and its comprehensive questions were typed on separate 8-1/2" x 11" sheets of white paper.

PROCEDURE

Vocabulary Presentation

The procedure for each story in the series commenced with the presentation of the new words introduced in that story. The words were presented individually on the cards, and S was asked to pronounce them. A correct response to a word-stimulus was eventually reinforced with a mid-value token. After a correct response to a word, the card was dropped from the group of cards yet to be presented. The S was instructed to indicate words whose meaning he did not know and this information was provided in such cases.

When an incorrect response to a word stimulus occurred, or when S gave no response, the instructional technician gave the correct response. The S then repeated the word while looking at the stimulus word. However, the word card involved was returned to the group of cards still to be presented. A card was not dropped from the group until it was read correctly without prompting. After an error on a word stimulus, only a low-value token was given for finally reading the word correctly without prompting. All tokens were delivered at the end of each phase rather than after each reading response.

Oral Reading

Upon completion of the vocabulary materials, each paragraph was individually presented to S in the order in which the paragraph occurred in the story. When correct reading responses were made to each word in the paragraph, a high-value token was given upon completion of the paragraph. When a paragraph contained errors, S was corrected, and he repeated the word correctly while looking at the word. The paragraph was put aside, and when the other paragraphs had been completed, the paragraph containing errors was again presented. The paragraph was repeated until it was read correctly in its entirety, at which time a mid-value token was presented. When all paragraphs in a story had been completed correctly, the next phase of the training was begun.

When a few of the Ss had some difficulty in reading the words in sequential order, either failing to respond to one or more words in the sequence, or adding words for which no reading stimuli were presented, the instructional technician would point to each word in turn and S would then read them. It was generally possible to drop this additional procedure later without any loss in the control over Ss' performance.

Silent Reading and Comprehensive Questions

Following the oral reading, S was given the sheet containing the story. He was instructed to read the story silently and he was told that it was important to read to understand the story so that he could answer the questions which would be presented later. Four yellow tokens were delivered at the end of the silent reading of the story.

Upon completion of the story, S wrote his answers to the questions typed on a separate sheet and gave his answers to the instructional technician. For each correct answer, S received a high-value token. For an answer with a spelling error, S was reinforced with a mid-value token when he had corrected the answer. For incorrect answers S had to reread the appropriate paragraph, correct his answer, and he then received a low-value token.

Vocabulary Review

Some of the vocabulary words presented to S in the first phase of training were words he

already could read. Many others, however, were words that the procedure was set up to teach. The oral-reading-phase performance indicated the level of S's retention of the words he had learned and also provided further training trials on the words not already learned. A further assessment of S's retention of the words that he did not know in the vocabulary training was made after each 20 stories of the SRA materials had been read. This test of individually presented words, for each story, was started immediately after completion of the 20 stories and constituted fairly long-term retention.

This test was also used as a review for S, and further training on the words was given. When S could not read a word, or missed one, he was prompted and had to correctly repeat the name of the word while looking at the word. This word-card was then put aside and presented later, at which time S was reinforced with a low-value token if he read it correctly. If not, the procedure was repeated until a correct unprompted trial occurred.

Facilities

The space employed for the training was varied and expedient, whatever could be obtained in busy and fully used schools. Thus, any room was employed which allowed two chairs and a table for support of the materials. Thus, offices, classrooms, cafeteria rooms, and so on, that were at that time not in use, were included. Sometimes more than one training group (subject and instructional technician) would work in a room at the same time, with no adverse effects.

second semester seniors in 33 schools in different parts of the United States. A mean score of 44.18 for 1,021 cases was secured and on that basis it was concluded that high school seniors are not adequately informed in the area of soil and water conservation. In addition it was suggested that secondary school curricula should be modified to include more effective teaching in this area. It is suggested that Glidden's 1953 study represents the first scientifically developed research approach to determine the basic principles of any area of conservation. A basis is also suggested for looking at differences of opinion about the importance of certain principles between groups of experts and among representatives of various regions of the country.

Caldwell (1955) conducted a study to determine what earth science principles to include in science programs in the general education curriculum of the secondary school and to determine the relative importance of each of those principles. The three phases involved in the study were: (1) the compilation of source materials in earth science from 13 textbooks, 8 reference books, 4 bulletins, and 2 research investigations; (2) an analysis of the source materials for statements of tentative earth science principles; and (3) a determination of the relative importance of 344 earth science principles which were selected to be included in the science program of the secondary school. A validity check was accomplished by comparing the investigator's selected list of 344 principles with a list derived independently by an assistant in biology who used the same technique on 50 pages of source material. After the list was submitted to two science teachers to eliminate duplication, 332 of the original 344 principles remained. The principles were then evaluated on a 3-point scale (+3=highly desirable, +2=desirable, -2=undesirable) by five educators to determine their relative importance for inclusion in the science program of the secondary schools. The ratings given were tabulated and the sum of the ratings for each recorded. With 292 principles remaining, the items were arranged in a descending order of importance. The form of the rating scale proved to be useful in describing the relative importance of each item and also in eliminating objectionable statements of principles from the list.

Trainor (1964) conducted an investigation designed: "To determine science concepts that should be taught in the junior high school science program;" and "To determine the grade level at which these concepts should be developed." Procedures included:

1. Developing a list of 357 concepts relating to junior high school science based on relevant literature.
2. Employing a survey technique including:
    a. the use of a questionnaire designed to accommodate additions to the list; and
    b. an interview to obtain comments from each respondent on ways to improve the list.
3. Conducting a pilot study to refine the questionnaire involving 18 selected educators.
4. Selecting a stratified random sample of 52 public schools in Rhode Island.
5. Delivering questionnaires to the principal of each school in the sample for distribution to selected elementary, junior high, and senior high school teachers.
6. Analyzing the data and establishing a minimum cut-off point of 50% agreement to be used in developing a final list of science concepts.

A pilot study with 16 of the 18 respondents completing questionnaires resulted in the development of a revised list of 285 concepts from the original list. Of the 164 questionnaires delivered to the randomly selected panel, 130 (79.3%) were completed and returned. These data indicated that 62 of the 285 concepts were to be developed in the junior high school. Further analysis indicated that at least 50% of the secondary teachers agreed that 190 concepts should be taught in junior high school. Only five concepts were indicated by Sixth Grade teachers as being appropriate for junior high school use but 75 concepts were indicated for use in elementary school and three in senior high school. It was concluded that the junior high school science curriculum needs to be re-examined and modified and that there is inadequate preparation of elementary teachers in science.

CONSERVATION EDUCATION CURRICULUM DEVELOPMENT

Early investigations in conservation education dealt with inventorying the extent of conservation teaching throughout the United States or in given geographic areas. Both attitudes and information possessed by students were of concern to Sherman (1950). The kinds of information being taught, the sources of conservation understandings, and the extent of conservation teaching were topics studied by Donnelly (1957), Hone (1959), Graff (1962), and Kenyon (1965). Sherman (1950) conducted a study designed to "...compare the conservation attitudes possessed by

III
RESULTS AND CONCLUSIONS

BEHAVIORAL DATA

Of the 36 experimental Ss who began the program, 32 were still participating when the study was ended. Four Ss left the program for reasons extraneous to the program—one S was not reinstated following suspension from school, and 3 Ss were sent to the reformatory. Those Ss who remained in the study participated in the program with excellent cooperation, attention, and work behaviors as the behavioral results show.

The total number of words each child read during the training was tabulated. This included the words he read when singly presented on cards, the words he read aloud in the paragraph reading phase, the words he read in the silent reading phase, the words read in the comprehensive questions, as well as the words read in the vocabulary review. During the mean 40.2 hours of participation, the mean number of word reading responses made by the Ss was 78,505. The cumulative record of the words read over the period of training for the two groups of Ss (those enrolled in regular classes and those enrolled in EMR classes) is depicted in Figure 1. It should be noted that one child began the experimental training late and was suspended in the middle of the program, and thus received by far the briefest training of all the Ss: 25 hours. One other child received only 32-1/2 hours of training and several others only 37-1/2 hours, which is close to the mean of 40.2 hours.

As was found earlier (Staats & Butterfield, 1965; Staats, Minke, Goodwin, & Landeen, 1967), the reading rate tends to accelerate across the period of training, even as the materials grow more difficult. This effect is most noticeable for the EMR group. Those Ss enrolled in regular classes show a slight but consistently more rapid rate over those Ss enrolled in EMR classes. This was due in large part to the greater amount of time spent by EMR children in learning of the individual vocabulary words. That is, more words were presented during this phase which were initially unknown to EMR children (a mean of 817.45 words were unknown to EMR Ss when first presented to them, as compared to a mean of 679.24 words unknown to regular Ss). Support for this contention was contributed by an analysis of covariance on the average reading rates of the Ss in these two groups, using the SRA pretest scores as the covariate. When these two groups were statistically equated in terms of the number of words initially unknown, no significant differences with respect to rate were obtained. The fact that there is no decrease in reading rate for either group, even at the end of the program, indicates that there was no decrease in the effectiveness of the reinforcer system, even after 50 hours of reading training. This result must be noted in view of these children's poor classroom performance.

Records were kept of the number of words the children missed on first presentation, the number of these words which were then later missed in the oral reading of the paragraphs, as well as the number of words originally missed that the child could not read on the review test presented at the later time. Ss from the EMR classes missed an average of 817.45 on the first presentation, retained 76.72% of these words after training when they were presented immediately following that training in the oral reading phase of the lesson, and retained 57.83% on the vocabulary review test presented a week or so later. Ss from regular classes, on the other hand, missed only 679.24 words on first presentation, retained 84.21% of these words on the short-term oral paragraph-reading retention measure, and retained 60.53% of these words on the vocabulary review test. To determine if the differences in retention obtained between EMR Ss and Ss from regular classes were significant,

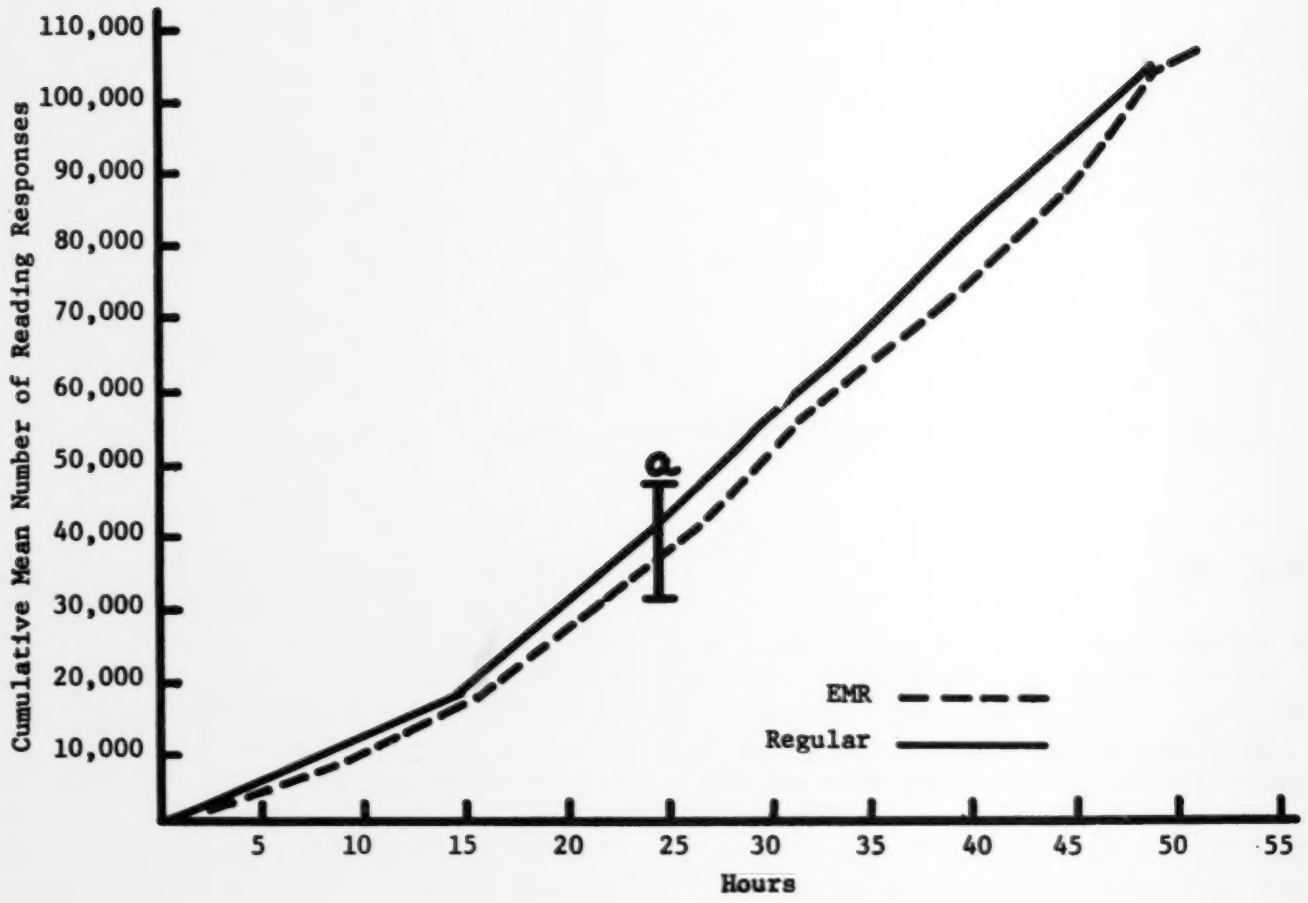


Figure 1. Cumulative number of reading responses as a function of the time in reading training for Ss enrolled in EMR and regular classes. The letter (a) indicates the last plotting point in which all Ss are represented.

both the short- and the long-term retention measures were computed for each S through Lesson 125 and an analysis of covariance was performed on these scores, with the SRA pretest score for each S serving as the covariate. (Four Ss did not complete the materials through Lesson 125 and were omitted from these analyses.) This analysis indicated that when the number of words initially unknown for each of the two groups was statistically equated, no significant differences in retention were obtained. Four Ss were eliminated from this analysis, since they did not complete 125 lessons during the course of the experiment.

The same analysis of covariance was performed on the number of learning trials required for words initially missed during the individual word phrase before S met the criterion of one correct unprompted trial and on the percentage of questions answered correctly during the comprehension phase of the lessons. EMR Ss required 1.33 learning trials per missed word in the individual word phase, as compared to 1.19 learning trials required by Ss in regular classes. This difference was significant at the .05 level. The percentage of comprehension questions answered correctly in the first 125 lessons was 94.40 for the Ss in the EMR group and 98.08 for the Ss in the regular classes ($p < .001$).

Evidence was presented in the earlier article that the 100-word test was a valid indicator of the Ss' performance with the SRA materials. Additional evidence was provided in this study in that the 100-word pretest score correlated .90 with the proportion of words correct when first presented in the individual word phase, .783 with the percentage of paragraphs read without error the first time presented in the oral reading phase, .71 with the proportion of words initially correct in the oral reading phase, and .57 with the reading rate of the subjects as indexed by the total number of reading response measures.

In the original study by Staats and Butterfield (1965), there seemed to be some evidence that as the training progressed the subject missed fewer of the words on first presentation. This type of evidence would tend to indicate that the subject had been learning to sound out new words as a function of the

training. Additional evidence (Staats, 1968) has shown that subjects <u>can</u> learn syllable reading units from training on whole word reading tasks. To test the possibility that improvement in sounding out new words occurred in the present study, the ratio of words missed the first time to the total presented was computed for each 20 stories presented. The results for the two groups are shown in Figure 2. The later points on each curve do not include all the subjects in the group, since some <u>S</u>s did not read as many stories as others did. In any event, there does not appear to be evidence that the ratio decreased, but rather that it was constant throughout the training except for an increase in the ratio from the 1.2 grade level to the 1.4 grade level of the reading material. These findings replicate those of the earlier study (Staats, Minke, Goodwin, & Landeen, 1967). The question of whether phonetic reading skills emerge from this type of training may be considered to remain unanswered. More direct tests will probably be necessary.

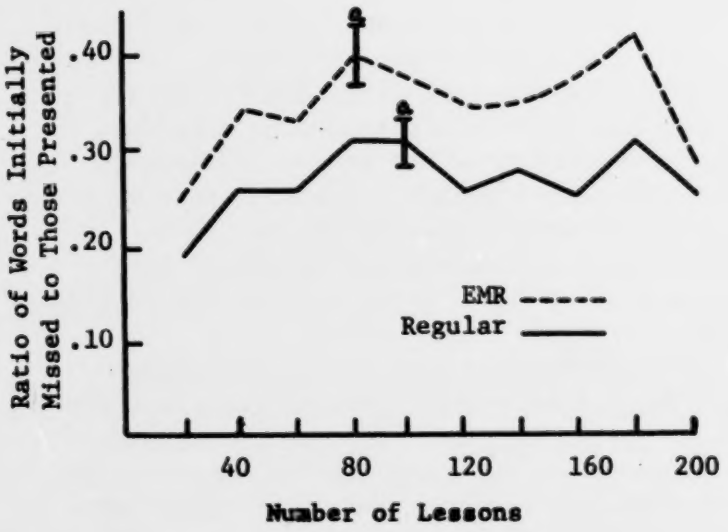


Figure 2. The ratio of words missed upon initial presentation to the total number of words presented in the Individual Word Phase as a function of the number of SRA stories read for the two classes. The letter (a) on each curve represents the last point in which all <u>S</u>s in that group are represented.

One of the important aspects of the procedures involves the ratio of reinforcement for the reading responses. The procedures were designed to progressively reduce the amount of reinforcement given per reading response as the training progressed or conversely to require more reading responses per unit of reinforcement. As was indicated in the earlier studies, demonstration that this is possible in a long-term training program has a number of important implications. Such a demonstration "is in part an answer to the question whether the use of extrinsic reinforcers in training will produce a child who is dependent upon these reinforcers" (Staats & Butterfield, 1965, p. 941).

Figure 3 supports the earlier demonstrations by showing that the ratio of the amount of reinforcement earned divided by the number of words read decreases as a function of number of training sessions. This result is especially interesting in view of the acceleration of rate of reading response shown in Figure 1. That is, in the period of time during which the reinforcement per reading response is cut in quarter, the rate of reading increases. Aside from its theoretical and practical implications for child learning, it is interesting to note that the result would be expected directly from basic laboratory research with lower organisms on the effects of reinforcement schedules. Thus, it is important to see that the basic principles hold even in the context of such a complex and uniquely human behavior as reading.

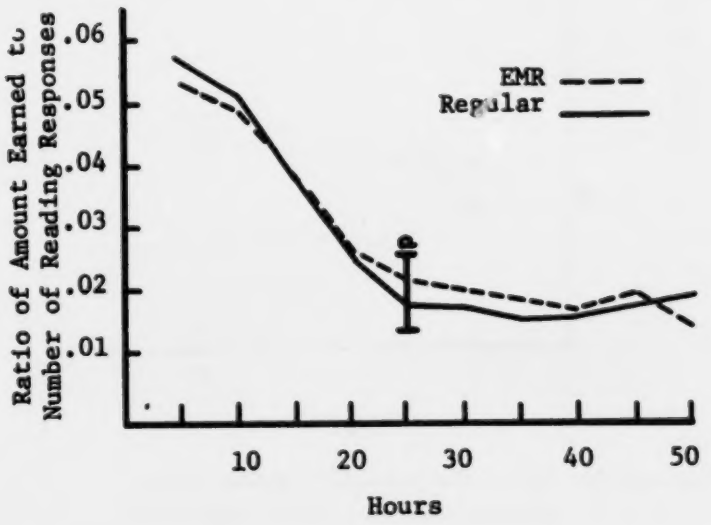


Figure 3. Ratio of the monetary value of the tokens received divided by the number of reading responses made as a function of time in reading. The letter (a) represents the last point in which all <u>S</u>s are represented, and it also indicates the time block in which the bonus was introduced for most <u>S</u>s.

Two types of grades were obtained on both experimental and control Ss: scholastic grades and deportment grades. For purposes of analysis, these grades were segregated with respect to whether they were given in academic or non-academic subjects. While no clear-cut findings were obtained with respect to deportment grades, an analysis of variance indicated that during the course of the experiment scholastic grades in academic subjects increased significantly for experimental Ss as opposed to control Ss from first semester to second semester (See Figure 4). The major part of the study was conducted during the second semester. Ss from one school began participation in the experiment the second week in November, one school began the first week in December, and one school did not commence the program until the first week in January, with the semester ending the second week of the same month. All three schools continued with the program until at least the third week in May, resulting in a minimum of 4 months of training taking place in the second semester for all Ss.

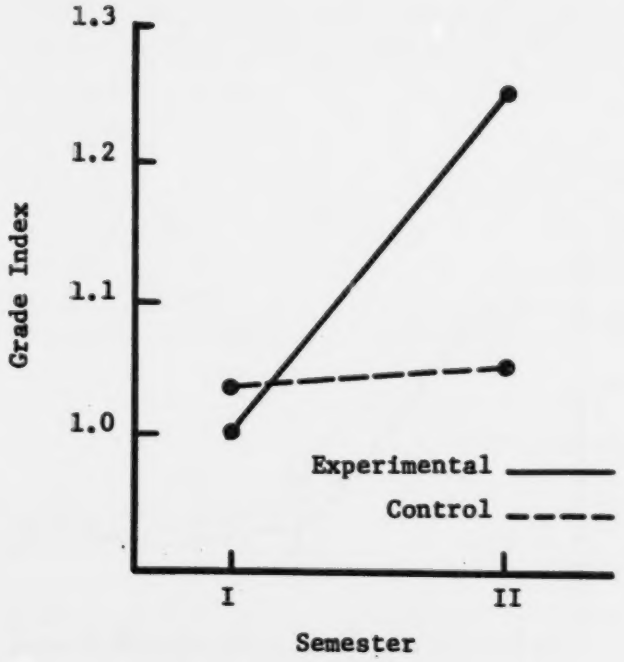


Figure 4. Mean scholastic grade point average in academic subjects for Experimental and Control Ss across the two semesters of the school year. An analysis of variance indicated that this treatment by semester interaction was significant ($p < .01$).

Attendance data were also obtained for both experimental and control groups. The Milwaukee School System records of student absences made it possible to classify cause of absence according to illness, suspension from school, or truancy. The proportion of total absences for the school year occurring after the onset of the experiment was determined for both the experimental and control groups under each of these three categories. The proportion of absences due to truancy during the experimental period was significantly less for the experimental group than for the control group ($p < .001$). On the other hand, the proportions of absences due to the other two reasons were quite similar and were not significantly different for the two groups of subjects.

The statistic employed to analyze the experimental and control group differences in grades and attendance was an analysis of covariance in the type of design to be described below.

TEST DATA

The design of the study with respect to experimental and control comparisons included three independent variables. The first was that of training versus no training, the experimental and control group comparison. The second variable was that of the EMR or regular class enrollment of the subjects. The third variable was that of the school in which the subjects were enrolled. Subjects came from three different schools in the ghetto area. An analysis of covariance was employed for the evaluation of the data and the appropriate pretest score was the covariate.

The dependent variables are the subjects' scores on the several tests at the end of the experimental training. The analysis shows whether or not there has been improvement on the posttests as a function of the three independent variables, with the covariate statistically equating each group with respect to the pretest score on each instrument.

The results of the analysis of covariance are shown in Table 1. As the Table indicates there were no significant results which stemmed from the subject's presence in one of the three schools. Moreover, no significant contribution was made by the classroom variable. That is, the children in the regular classes did not improve on the tests significantly more than the children in the classes for the educationally mentally retarded. Under the reinforcement training conditions, which guaranteed the attentional and work behaviors of the subjects, their respective learning abilities did not contribute a significant difference.

Table 1
Analysis of Covariance on Pre- and Post-Test Measures

Source	df	100-word Test		Vocabulary		Reading		Verbal IQ		Nonverbal IQ	
		MS	F	MS	F	MS	F	MS	F	MS	F
Treatment (A)	1	694.20	151.63***	0.47	2.21	0.72	3.33*	0.79	2.54	8.92	4.40**
Class (B)	1	46.15	1.01	0.13	---	0.50	2.34	0.07	---	0.34	---
School (C)	2	59.01	1.29	0.12	---	0.11	---	0.13	---	0.02	---
A x B	1	60.72	1.32	0.09	---	0.00	---	0.02	---	0.01	---
A x C	2	43.01	---	0.46	2.16	0.32	1.47	0.20	---	4.23	2.09
B x C	2	7.54	---	0.18	---	0.14	---	1.16	3.72**	3.77	1.86
A x B x C	2	100.18	2.18	0.11	---	0.02	---	0.33	1.05	0.11	---
Within	51	45.93		0.21		0.22		0.31		2.03	

*$p < .10$ **$p < .05$ ***$p < .001$

On the other hand, the effects of the training on the experimental group's learning as compared to that of the control were clearly shown by their respective results on the 100-word reading test. That is, the experimental group could read a mean of 26.23 more of the items on the 100-word test after their training than they could before the training had commenced. For the control group the gain was only 5.66 words during the same period of time. An analysis of covariance was conducted to test the statistical reliability of the difference, with the pre-experimental 100-word test results as the covariate. The difference was significant at far beyond the .001 level.

As mentioned earlier, one of the major problems encountered in the previous study was the failure of treatment differences to be reflected on standardized achievement tests. It was suggested at that time that such differences might have been obtained had a lower form of the test been used. To test this hypothesis, both the achievement tests and the intelligence test employed in this study were normed on a school population younger than the Ss employed in the experiment. The Metropolitan Achievement Tests employed were designed for use with Grades 3 and 4 rather than Grades 7 and 8. Experimental Ss demonstrated relative means gains over control group Ss on both the vocabulary and reading portions of this test. Thus, experimental group Ss demonstrated gains of .20 of a year as opposed to a gain of .06 of 1 year for the control Ss on the vocabulary test, and .37 of a year on the reading portion, as opposed to .20 of a year for the control group. An analysis of covariance was performed on both of these test scores, using the obtained pre-test scores as the covariates, thus statistically equating both groups with respect to their pre-experimental achievement scores. The difference between groups on reading portion of the Metropolitan Achievement Test was significant at the .074 level. The vocabulary test results would be expected by chance with a probability of .143.

An analysis of covariance on both the verbal and non-verbal portions of the lower form of the Lorge-Thorndike intelligence tests also indicated a significant difference due to treatment. When the raw test scores were converted into age equivalent scores, the experimental group demonstrated a gain of .50 of a year on the verbal portion of the test and a gain of 1.35 years on the non-verbal portion, as compared to a verbal gain of .26 and a non-verbal gain of .70 for the control group. The former difference would be expected by chance with a probability of .117, while the difference between the experimental and control group on the non-verbal portion of the test was significant at the .05 level ($p < .041$). It should be noted that the age equivalent scores were used in these analyses rather than IQ scores since the test was normed with younger children, which invalidates the use of IQ scores.

Table 2, cont.

Doctoral Dissertations
Hanselman, D. L. Interdepartmental teaching of conservation at The Ohio State University. Natural Resources Institute, The Ohio State University, Columbus, Ohio, 1963.
Visher, H. H. A determination of conservation principles and concepts desirable for use in the secondary schools. University of Indiana, Evansville, Indiana, 1960.
White, R. C. A study associating selected conservation understandings with available community resources for grades four through twelve. University of Montana, Missoula, Montana, 1967.
Yambert, P. A. A principle and concept perspective of conservation. University of Michigan, Ann Arbor, Michigan, 1961.

Establishing Concept Credibility

The initial list of 89 concepts was sent to the Wisconsin Panel along with a letter indicating the nature of the participation desired of them; each panel member was to accept or rewrite the concepts listed and to make additions wherever necessary. Each panel member was also contacted so that a time for a personal interview could be scheduled. During this interview the completed survey instrument was picked up and each scholar was requested to respond to the question: "What should a student know about environmental management?" A standard form of questioning was used and all relevant comments made by the scholar were recorded by the interviewer.

Revised Concept List No. 2

A revised list of 157 concepts based on the results of the survey instrument and interviews was formulated into an ihstrument utilizing a modified Blanchet scale technique. This scale provided the respondent with an opportunity to make one of four acceptable choices indicating descending order of importance or one unacceptable choice [Essential (+5), Highly Desirable (+4), Desirable (+3), Satisfactory (+2), Unacceptable (-5)] relative to each concept. In addition each panel member was invited to judge individual item credibility.

Revised Concept List No. 3

A third list of 128 concepts, based upon the reactions to List No. 2, was formulated again utilizing the Blanchet scale. The revised list was sent to the 699 member National panel along with a cover letter explaining the nature of the participation desired for evaluation in terms of credibility and degree of acceptability.

TREATMENT OF DATA

Wisconsin Phase

The written comments and additions made by the Wisconsin Panel on Concept List No. 1 were recorded and used in revising the initial list.

Data consisting of the reactions of the Wisconsin Panel to List No. 2 (Blanchet scale) were tabulated using the Reciprocal Averages Program (RAVE) (Baker, 1968) and the frequencies of response for each of the five categories were determined. A 75% level of total acceptability was established as a criterion to exempt a concept from further revision or exclusion.

National Phase

The List No. 3 data were tabulated using RAVE and the frequencies for each of five categories of response were determined. A 90% level of acceptability was established as requisite for concepts to be included in the final list. Therefore, any concept receiving 10% or more unacceptable responses was eliminated from the list and the reasons for its exclusion were analyzed.

GPO 819-216-4

IV
DISCUSSION

The subjects in the study may be considered to be rather typical of the urban, ghetto, culturally disadvantaged child population. Education has found children from this group to constitute its most difficult problem. Traditional educational practices have failed markedly with such children. These children to a large extent do not attend or work in the classroom and they do not learn. Furthermore, the subjects in the present study were selected to be among the most difficult educational and social problems in this difficult group. These were generally children who even relative to this group, did not attend well in class, who misbehaved, who had high absenteeism, who did not learn. One third of the subjects were considered to be educable mental retardates and were in special classes. Many of the children were severe behavior problems and had been in reformatories or detention homes. This is stressed because the results show that when such children are subjected to a suitably designed set of stimulus materials and procedures, utilizing a reinforcement system also suited to the children, the attention, participation, and work behaviors of the children become quite good. Children who did not learn in the typical classroom became hard workers and good learners when the materials and procedures were employed and when their behavior was rewarded by use of the token-reinforcement system.

There are several different types of data that show this result clearly. First, there is the observation of the children's behavior while in the training sessions. The typical behavior was one of good attention and rapid work. Objective records of the children's progress stipulates this clearly. That is, during the period of the experimental training the subjects covered a mean of 78,505 single-word reading responses. This finding is supplemented by the fact that the cumulative record of the rate of making reading responses shows a positive acceleration over the training period. Thus, although the reading material becomes steadily more difficult, the subjects made more rapid reading responses as the training progressed. This is objective evidence of the fact that the children's attentional and work behaviors were increasing in strength. It is interesting to note that the acceleration in rate of reading responses is shown for the retarded children as well as for the children in regular classes. The form of the curve suggests an early improvement in the constituent reading behaviors yielding an early acceleration followed by a linear curve indicating the behaviors were maintained in good strength in the face of increasingly more difficult material. Actually, the qualitative observations of the children's attentional and work behavior suggested that the token-reinforcement procedures had very strong effects. That is, the present children's work was even better than occurred in the previous study of Staats, Minke, Goodwin, & Landeen (1967). This study was conducted with Caucasian children from more culturally advantaged homes; as the whites commenced with higher level of reading skill, the task was a more difficult one for the present children.

Supplementary evidence of the efficacy of the token-reinforcement procedures for maintaining the participation of the experimental subjects is given by the comparison of absenteeism in school with the control group. The result shows that the experimental group had fewer truant absences than the control group, even though there were not significant differences between the groups in terms of absences for uncontrolled reasons such as illness or suspensions. These results suggest that the reward-value of the remedial training employing extrinsic reinforcement was strong enough to exert the more general effect of increasing school attendance. If 1/2 hour per day appears to exert some effect in this direction, one must

wonder what could be done with the employment of reinforcement principles in a more extensive manner for these children. One fact is irrefutable: learning trials are necessary for learning to occur. At the most basic level procedures have to be devised for these children to insure that they are present in school and more centrally that they attend and work when they are present. Traditional procedures do not provide these guarantees. The present reinforcement procedures appear, in principle at least, to strengthen the basic behaviors essential to academic learning. Moreover, this can be done without committing the children to institutions.

In addition to the evidence that the token-reinforcement procedures dramatically reversed the poor attentional and work behaviors of the experimental subjects, the records made of each child's responses indicated that the children were also learning a good deal of new material. Thus, for example, the children missed a mean total of 726.75 of the various words that were singly presented to them in the word-learning phase of the procedures. The percentage of these words read correctly when they later occurred in the context of a paragraph was 81.07%. Thus, there was good retention of the words learned when presented a short time later and in context. The measure of long-term retention when the words were individually presented was 59.40%. It seems a rather remarkable performance for these subjects, one-third of whom were considered to be retarded, to retain 60% of the material on which they received learning trials.

The first study utilizing the present materials, procedures, and reinforcement system was conducted with a culturally deprived juvenile delinquent (Staats & Butterfield, 1965). It is interesting to note the close generality and replicability of the methodology. Thus, all of the 50 subjects in the second two studies (Staats, Minke, Goodwin, & Landeen, 1967, and the present study) have responded in a like manner in terms of attention and work behaviors. Moreover, the learning proceeds in a standard manner with the differences coming only from the children's beginning level of reading skill. In the first study the subject missed a mean total of 761 words, had a mean short term retention of 77%, and a mean long term retention of 57% (Staats & Butterfield, 1965). The relevant figures for the present EMR subjects were 817 words learned, 77% short-term retention, and 58% long-term retention. For the regular Ss the figures were 679 words, 84%, and 61%. It is interesting to note that the study of Staats, Minke, Goodwin, and Landeen (1967) dealt with more skilled children. They missed a mean total of 593.5 words, had a mean short-term retention of 86%, and a mean long-term retention of 68.4%. The long-term retention of the present subjects seemed very high in view of the fact that the subjects in the Staats, Minke, Goodwin, and Landeen study had greater reading skill at the beginning of the training. That is, the other children had an initial mean score of 63.8% on the 100-word sample test versus the mean score of 38.81% for the present subjects. It appears, thus, that when learning is conducted to the criterion of only one spontaneous reading of a word, and the word is later read in context, generally effective learning and retention are produced.

It is also interesting to compare the children in the regular classes in the present study to the children in classes for the retarded. The differences in rate of reading and retention between these two groups were found to be a function of the number of new words the children in each group had to learn. That is, when the number of words initially unknown for each of the two groups was statistically equated by means of an analysis of covariance reading rate and retention, differences were found to be non-significant. Even when this concommitant variable was controlled, however, significant differences were obtained between those children in regular classes and those in EMR classes on the number of learning trials required to master an initially unknown word and on the percentage of comprehension questions which were answered correctly following completion of the stories.

Although these findings are only suggestive, a procedure whereby child learning involving a complex cognitive task can be observed and measured along a number of parameters across an extended period would seem to be particularly appropriate to the study of how retarded children learn and the nature and extent of their actual behavioral deficits. As has been discussed elsewhere (see Staats, 1968a) such information may then serve as the basis for making a detailed analysis of the learning involved and should ultimately result in procedures designed to maximize that learning.

The present study contained additional means of assessing the learning of the children in the training procedures. The most direct test of the comparative learning of the experimental and control groups involved comparing them upon the materials on which the experimental group received training (the 100-word test). It must be noted that the 100-word test was a sample of all several thousand different words which were presented to the experimental group children in the reading materials. The experimental

group could read a mean of 38.81% of the sample of words prior to training. The control group could read a mean of 37.56% of the sample at that time. After the training program the experimental group read 65.34% of the sample of words. During this same period of time, however, the control group made a much lesser advance. They advanced to a mean percentage of 43.22 on the 100-word reading sample. The difference in learning and retention was significant at much beyond the .001 level. Thus the experimental group improved from a mean of 38.81 words out of 100 to 65.34. This is a 69% gain in these children's word reading ability. The control group in advancing from 37.56 to 43.72 showed only a 15% gain. The experimental training thus had a large and positive effect upon the word reading skill.

It is interesting to note evidence that the 100-word test was a representative sample of all the words in the SRA materials and a good indicator of performance in the reading task. That is, the proportion of total words the child could read on first presentation was correlated .90 with the 100-word test performance, the proportion of paragraphs read correctly initially was correlated .78, the proportion of words on the short-term retention measure was correlated .71, and the long-term retention measure was correlated .74, at better than the .001 level. On the other hand, there was a less close correlation of .25 between the 100-word test performance and reading comprehension (answering the SRA questions), the correlation being non-significant.

In addition to these direct measures, the reading and intelligence tests were also given to the experimental and control groups to assess any effect of the SRA reading materials upon standard test performance. It is interesting to examine the types of items on these tests in terms of the training given the experimental subjects and in terms of experimental control comparisons. First, the test that tapped most closely skills on which the experimental subjects had been trained was the reading part of the Metropolitan Reading Achievement Test. That is, although there was probably little overlap between the words the experimental subjects learned and the reading test, the other skills involved in reading (see Staats, 1968a) would be expected to generalize. This expectation was born out in the results; the experimental group gained significantly more (at the .07 level) from pre- to post-test than did the control group. What is called the vocabulary part of the test actually involves a great deal more than reading words, however. That is, for example, the items require that the subject be able to place words in the correct superordinate class in answering the multiple choice items. Success on such items depends to a large extent upon the complexity of the child's language learning in various spheres, not just the skills involved in the present training. The 20th item, for example, states the following: things that are separate are . . . alike, united, apart, or together. Thus, on this task although the difference was in favor of the experimental group, the significance level was only .14.

On the intelligence test there were similar results. The non-verbal part of the intelligence test demanded good attentional responses and discrimination, some reading of letters, and a good deal of number reading. This part of the test thus actually concerned skills more similar to those on which the experimental group subjects were trained than did the verbal part of the intelligence test. It is thus interesting to note that the difference between the experimental and control groups was significant on the non-verbal part of the Lorge-Thorndike at the .04 level. The verbal part of the intelligence test, on the other hand, was like the vocabulary part of the reading achievement test. The verbal part involved testing of the child's level and complexity of language organization. The difference between the experimental and control groups was still in the right direction even on these items. However, the level of significance was only .12.

Thus, all of the test results are in the direction favoring the experimental group, although only two attain or approximate an acceptable level of statistical significance. The fact that the experimental subjects had an advantage over the control group on all measures heightens our confidence in the reliability of the effect over and above the verification given by any one measure. In conjunction with the other results and with an analysis of the tasks involved on the various parts of the tests, it is suggested that there is evidence that the training which the experimental subjects received, and the control subjects did not, had general effects. That is, even on materials very much different from those on which the subjects received training, the experimental subjects displayed somewhat of an advantage. The advantage did not approach the magnitude shown when the children were tested on materials on which the experimental subjects had received training, however. The combined results nevertheless suggest that the SRA materials when adapted into the present stimulus presentation form constitute a training program which produces reading learning and cognitive development in children with learning deficits.

It should be indicated that the above results are suggestive that the experimental training may be reflected on standardized instruments. However, it cannot be concluded that the period of training of the study was sufficient to remediate in a complete manner the experimental subjects reading deficits, or to improve to any great extent their general deficits of language and cognitive skills. Although the reinforcement procedures were highly effective in obtaining attentional, participation, and work behaviors from the children so that they would learn well, it must be concluded that more extensive training would be necessary to effect more extensive cognitive learning. A one-semester program of daily half-hour sessions is not enough. However, it may be suggested that the present study demonstrates in principle the efficacy of the application of the learning principles. If these positive effects can be obtained in 40 or so hours of training the implication is that very substantial gains could be made with longer periods of training. The results indicate these types of children should be involved in a program which insured many more learning trials than were involved in the present study—employing the present types of procedures for insuring attention and learning. The program would have to be for a greater length of time per day, and for a greater length of time than one semester. Thus, a program should be conducted to expand the training in these directions, employing the types of reading materials used in this study. The present results plus preceding ones in the first author's long-term study of reading give strong support to the learning principles employed as well as the reinforcement procedures.

One of the primary purposes of the study was to investigate the possibility of using black adults as instructional technicians who would actually conduct the teaching. The method of having a brief training period for the instructional technicians, with supervision by a trained teacher, proved to be quite effective. It was interesting especially to note that the formerly unemployed full-time instructional technicians performed very well in the program. The work was satisfying to them, the teaching was advantageous and rewarding to their pupils, and so on.

It may be suggested that the present program may be taken as a model of organization which has general merit. That is, the model would consist of teams of supervisory teachers and instructional technicians. This would raise the professional level of the teacher as well as that of the instructional technicians. Such a model would also provide very important and useful employment to large numbers of black adults in a situation which would introduce them to new values of various kinds, especially the value of education. It would be expected that adults who were so employed would raise children more oriented towards and prepared for achieving educational success. It would also be expected that if there were such positions of employment, there would be greater incentive for children and young adults to gain the educational skills (such as reading) to qualify for the instructional-technician positions. Moreover, this is a model that could be put into effect now. That is, many black adults, presently unemployed, could be trained to read themselves (if this were a problem) and could then be employed as instructional technicians in programs to remediate the reading deficits of educationally retarded ghetto children. The same model would be appropriate also for other academic areas, such as arithmetic and mathematics, and so on. Larger field studies should be made of these possibilities.

That is not to say that there would not be a cost involved in such a long-term program. Although the reinforcers for the children are not an expensive item—a mean of $21.34 for a mean of 40.2 hours of participation—the hiring of full-time instructional technicians would involve another cost. However, when one considers the cost to society of unemployed adults and of their children who will later be unemployed because of academic failure, the cost of the instructional technician program would actually be less.

It may be added that the other instructional technicians, the high school seniors and the adult volunteers, also performed effectively in the program. In only two cases was there reason to replace an instructional technician for not being able to adhere to the character of the procedures. In general, however, it can be expected that use of full-time instructional technicians would entail more advantages than the use of part-time personnel. The training time per instructional technician would be reduced and, more importantly, it could be expected that the person employed full time over a long period would develop additional skill in teaching children.

Finally, it may be suggested that the study supports the validity of laboratory-derived principles on the human level in the context of the acquisition of complex cognitive skills. In so doing generality of the basic principles is shown. Moreover, this adds to one's confidence in the learning conception of human behavior. As has been suggested (Staats, 1968a, 1968b) psychology must begin the study of functional repertoires of human behavior,

employing its basic principles and experimental procedures. Such investigations of functional repertoires have more widely commenced in clinical psychology. Educational, child, and social psychology are not as advanced in this respect. This is clear, for example, in educational psychology, where most studies still involve experimental tasks which are labeled as though they represented some universe of functional human behaviors but the universe is never made explicit and the study remains academic in the pejorative sense of the word. How the findings of the majority of studies labeled problem solving, concept formation, verbal learning, and so on, extend to real samples of such cognitive skills is never made clear. It is time that representative samples of actual human behavior be systematically studied using the basic principles of learning. The present study demonstrates the possibilities for doing this. The project of which this study is a part more generally demonstrates that experimental methods and analyses of complex learning tasks may be employed in studying and treating ever more representative samples of human behavior in the continued extension and elaboration of the learning theory. In the extended project in the study of reading acquisition as a type of cognitive learning, the research has moved from basic research in the laboratory-controlled situation to this field study sort of experiment. The present study may thus be valued/as part of this strategy of psychological theory construction, as well as for contributing to the solution of an important problem of human behavior.

REFERENCES

Dolch, D. W. Teaching primary reading. Champaign, Ill.: Garrard Press, 1950.

Staats, A. W. Learning, language, and cognition. New York: Holt, Rinehart, and Winston, 1968(a).

Staats, A. W. Categories and underlying mental processes, or representative behavior samples and S-R analyses: Opposing heuristic strategies. Ontario Journal of Educational Research, 1968, 10, 187-201.(b)

Staats, A. W. A general apparatus for the investigation of complex learning in children. Behavior Research and Therapy, 1968, 6, 45-50.(c)

Staats, A. W. Social behaviorism, human motivation, and the conditioning therapies. In B. Maher (Ed.), Progress in experimental personality research. New York: Academic Press, 1970.

Staats, A. W. and Butterfield, W. H. Treatment of nonreading in a culturally deprived juvenile delinquent: An application of learning principles. Child Development, 1965, 36, 925-942.

Staats, A. W., Finley, J. R., Minke, K. A., and Wolf, M. M. Reinforcement variables in the control of unit reading responses. J. exp. anal. Behav., 1964, 7, 139-149.

Staats, A. W., Finley, J. R., Minke, K. A., Wolf, M. M., and Brooks, L. O. A reinforcer system and experimental procedure for the laboratory study of reading acquisition. Child Development, 1964, 35, 209-231.

Staats, A. W., Minke, K. A., Goodwin, W., and Landeen, J. Cognitive behavior modification: "Motivated learning" with subprofessional therapy-technicians. Behav. Res. Therapy, 1967, 5, 283-299.

Staats, A. W., Staats, C. K., Schutz, R. E., and Wolf, M. M. The conditioning of textual responses using "extrinsic" reinforcers. Journal of the Experimental Analysis of Behavior, 1962, 5, 33-40.

Ullman, L. P. & Krasner, L. (Eds.). Case Studies In Behavior Modification. New York: Holt, Rinehart and Winston, 1965.

Wolf, M. M., Giles, D. K., and Hall, R. V., Experiment with token reinforcement in a remedial classroom. Behavior, Research and Therapy, 1968, 6, 51-64.

GPO 818-218-3